

RECEIVED

AUG 10 2000

TECH CENTER 1600/2000

SEQUENCE LISTING



(1) GENERAL INFORMATION

- (i) APPLICANT: Rine, Jasper D.
Boyartchuk, Victor L.
Ashby, Matthew N.
- (ii) TITLE OF THE INVENTION: AFC1 AND RCE1: ISOPRENYLATED CAAX
PROCESSING ENZYMES
- (iii) NUMBER OF SEQUENCES: 4
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Fish & Richardson P.C.
 - (B) STREET: 2200 Sand Hill road, suite 100
 - (C) CITY: Menlo Park
 - (D) STATE: CA
 - (E) COUNTRY: USA
 - (F) ZIP: 94025
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 09/167,132
 - (B) FILING DATE: 05-OCT-1998
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/902,774
 - (B) FILING DATE: 30-JUL-1997
 - (A) APPLICATION NUMBER: 60/023,491
 - (B) FILING DATE: 07-AUG-1996
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Suyat, Reginald J.
 - (B) REGISTRATION NUMBER: 28,172
 - (C) REFERENCE/DOCKET NUMBER: 09272-006003
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 650/322-5070
 - (B) TELEFAX: 650/854-0875

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1825 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: Genomic DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ACCTACCTTT TTTTCTATCT TCAACAACGA AACGCCTTAC ACACACACAC ACATACATCT

60

ACATACATAC	ATACAAATAT	ACATATATGT	AAACTTGTAT	ATTCATTCCCT	ATTAACCAAA	120
AAGAGGCAAT	TAAACTTTTC	CCTCTTTTTC	TACGTCATTT	ACTCAAAAAC	TCTAATTCCT	180
TCGTCTCTGT	TCTGCCATTT	TCTCCAGAAA	AAAATCGACG	GGAAATAAAA	AAAAAAGAC	240
AACGAACAAG	AGAAAAAGTT	CGCGAATTAT	AAACCACTTC	TATAATTAAC	AGGAAAAGGA	300
AGGAAAAAAA	AGGAGGAAAA	GAAAACTGCA	GGCCTTTTAT	CATGTTTGAT	CTTAAGACGA	360
TTCTCGACCA	TCCTAATATC	CCGTGGAAAT	TAATCATTTC	TGGGTTCTCG	ATTGCCCAAT	420
TTTCTTTTGA	ATCTTACTTG	ACGTACAGAC	AGTACCAGAA	GCTATCTGAA	ACAAAGTTGC	480
CACCTGTGCT	GGAAGACGAA	ATTGATGATG	AAACTTTTCA	TAAATCAAGG	AACTACTCCC	540
GGGCAAGGC	CAAGTTCTCC	ATTTTCGGTG	ACGTCATATA	CCTAGCCCAA	AAGCTAGTTT	600
TCATCAAATA	CGACCTCTTC	CCTAAAATCT	GGCACATGGC	CGTTTCTTTA	TTGAATGCAG	660
TCCTGCCAGT	CAGATTTTCAT	ATGGTCTCCA	CTGTGCGACA	GAGTTTTATG	CTTCTTGGGT	720
CTCTTATCCA	GTTTGTCTAC	CTTGGTTGAT	TTGCCACTCT	CTTACTATAG	CCATTTTGTC	780
CTGGAAGAAA	AATTTGGTTT	CAATAAATTG	ACCGTCCAAC	TATGGATCAC	CGATATGATC	840
AAGAGTCTGA	CTTTGGCGTA	TGCTATTGGT	GGCCCAATCC	TTTACCTGTT	CCTTAAGATC	900
TTTGATAAAT	TCCCTACTGA	TTTCCTTTGG	TACATTATGG	TCTTCTTGTT	CGTTGTCCAA	960
ATCTTAGCCA	TGACAATCAT	TCCAGTCTTC	ATCATGCCCA	TGTTTAATAA	GTTCACTCCA	1020
TTGGAGGACG	GTGAACTGAA	AAAATCTATT	GAAAGTTTGG	CCGATAGAGT	TGGGTTCCCT	1080
CTAGATAAGA	TTTTTGTCAT	TGACGGCTCA	AAAAGATCTT	CTCATTCAAA	CGCATATTTT	1140
ACAGGTTTGC	CATTACCTC	CAAGAGAAAT	GTTTTGTTCG	ACACTTTAGT	GAACAGTAAT	1200
TCTACTGATG	AAATTACGGC	TGTTTTGGCC	CATGAAATCG	GTCAGTGGCA	AAAAAACCCAC	1260
ATCGTTAATA	TGGTCATCTT	TAGTCAATTG	CACACCTTCC	TCATTTTCTC	CCTTTTCACC	1320
AGCATCTACA	GAAATACATC	ATTTTACAA	ACCTTCGGCT	TTTTCTTAGA	GAAGTCCACT	1380
GGCAGTTTTG	TTGATCCCGT	TATCACTAAG	GAATTCCTCC	TTATCATTGG	ATTTATGTTA	1440
TTTAACGACT	TATTAACCTC	ACTCGAATGT	GCCATGCAAT	TCGTGATGAG	TTTAATTTCC	1500
AGAACTCATG	AATATCAAGC	TGATGCTTAT	GCTAAAAAAT	TGGGCTACAA	GCAAAATCTA	1560
TGTAGGGCTC	TAATTGATCT	ACAAATCAAA	AACCTTTCCA	CCATGAATGT	AGATCCTCTG	1620
TATTCTAGAT	ATCATATTTC	CCATCCAAT	CTAGCTGAAA	GATCGACCGC	TCTAGACTAT	1680
GTTAGTGAAA	AGAAGAAAAA	CTAATCTATA	GAGTACACAT	ATTAGCATGT	ACCGTTAAAT	1740
TCAGCTTCGT	TATGTCTATA	TCTACATACA	TACACAGGTA	TCTACTATAA	GAATAAAGGA	1800
AAGAAAAAAT	AAACGATTAA	ACATT				1825

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2850 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TGAACTGTTG	ATGAACAAAG	AGAAGCTGAC	AAGCATCAAA	GCTTTGTACG	ATGATTTCCA	60
TTCAAAAATT	TGTGAATATG	AAACCAAGTT	CAACAAGAA	TTTCTTGAAT	TAAATGAGTT	120
ATATAATATG	AATAGGGGAG	ACCGTAGGCT	AAAGGAACTG	AAATTTACAG	ATTTTATTAC	180
TTCACAGCTG	TTTAACGATA	TCGAAAGCAT	TTGCCAATTG	AAAGTTAGTG	TTCACAACTT	240
ATCCAACATT	TTTAAAAAAC	AGGTCAGTAC	CCTAAAACAA	CACTCAAAGC	ACGCATTATC	300
TGAGGATTCA	ATATCGCACA	CAGGTAACGG	TAGTTCATCG	TCGCCCAGTT	CAGCGTCATT	360
AACGCCAGTA	ACTTCTTCAT	CCAAGAGTAG	TTTATTTTTA	CCTAGCGGTA	GCTCGTCTAC	420
TTCCCTGAAA	TTTACAGACC	AGATTGTTCA	TAAATGGGTT	AGGATTGCTC	CTTTACAGTA	480
CAAACGAGAC	ATTAATGTGA	ACTTGGAATT	TAATAAGGAC	ATTAAGGAAA	CTTTAATTCC	540
AAGTTTTGAA	AGCTGCCTAT	GTTGTAGGTT	TTATTGCGTT	CGAGTAATGA	TTAAATTGTA	600
AAACCATCTT	GGCGTAGCGA	AGATTGATAT	CCCTATTTCT	GTTAGGCAAG	TGACAAAATA	660
AAAAAACATT	AGAAAAAATT	CTCGTTACTT	TTCTTATAGA	TATAGATATA	TGTATGGTTT	720
GCTTATAGAT	GAAGGTATTT	ATCGCGTCCT	TTGTATTCCC	TATTATTAAT	AAAATTCCTT	780
TAAATGCAT	TTTCTGGTGC	TCTTTTGTG	CTTCTGTATT	TTTTTTTTTT	TGGACCACTG	840
GATGGAAAAC	CTTTGATGAT	TTTATTACCT	TTATTTTAAC	TTACTAAAAT	ATCGAGATTT	900
CAGGAACAAA	ACATAGAATT	TTCTTTGTCA	AGAAAAATAA	AACGAAATAA	ATTGATGCTT	960
TGACTACTGA	CTGTCTGTCA	TAGAGAGAAC	CAGAACAGCA	ATGCTACAAT	TCTCAACATT	1020
TCTAGTGCTC	CTATACATCT	CCATATCCTA	TGTGCTACCG	CTATATGCAA	CTTCACAACC	1080
AGAAGGGTCT	AAACGAGATA	ATCCTCGAAC	GATTAAATCT	CGCATGCAAA	AACTTACAAT	1140

TATGCTAATT	TCCAACCTTT	TTTTGGTGCC	TTTTTTACAA	TCTCAATTAT	CTAGTACCAC	1200
TTCACATATA	AGTTTCAAGG	ACGCATTTTT	AGGCTTAGGT	ATTATCCCAG	GTTATTACGC	1260
TGCATTGCCA	AACCCTTGGC	AATTCAGCCA	GTTCGTGAAA	GACTTAACGA	AATGTGTTGC	1320
GATGTTATTG	ACCTTATATT	GTGGACCCGT	TTTAGATTTT	GTATTATATC	ATTTATTAAA	1380
TCCAAAGAGC	TCTATACTTG	AAGATTTTTA	CCATGAATTC	CTGAATATTT	GGAGTTTCAG	1440
GAATTTTATA	TTTGACCCAA	TAAC TGAGGA	AATATTTTAC	ACGTCAATGC	TTTTGACTAC	1500
GTACTTAAAC	CTAATACCGC	ATTCGCAACT	AAGCTATCAA	CAGTTATTTT	GGCAACCATC	1560
GCTTTTTTTT	GGACTTGCGC	ACGCACACCA	TGCTTATGAG	CAATTACAGG	AAGGCTCCAT	1620
GACAACTGTT	TCCATTCTGC	TGACAACATG	CTTCCAAATT	TTATACACAA	CACTTTTTTG	1680
AGGGTTAACC	AAGTTTGTAT	TGCAATATCA	TGGGGTTTAA	TGGTCCTTCA	AGATTGAATT	1740
TACATTTTAC	ACCTTAGAGC	AAGAAAGCTG	GACGCATTTT	CAAATTGGTC	TCTCAATCTG	1800
GAATAAGTGC	TACTTCGCAC	TGCTGGTCCT	TGGATTAATA	TCCCTGAAGG	ATACCTTACA	1860
AACCTCTGGT	GGAACCTCTG	GTTATAGAAT	AACCCTTTAG	CCTTTTTTAC	GTACTTGTAT	1920
ACCGTTTAAA	ATTTCCTATG	TACTATAACC	TTTTTTTCACT	ACTATTATGG	AATTCTATCG	1980
AGCGACCGGG	CTTTTGTTAC	GGAAGAGTGA	AAAAATCGAG	TTTTGGTGTT	TTGGTGAAAG	2040
AATTTGGAGG	ACTATAAAGT	ACCTATACTT	TGTATTACGG	ACTCAATAAC	AAGTCGTTCC	2100
TGTCAGTGGT	ATTGAAGTTG	TCAGATCTAA	GAGTAGAGAG	AAGGTGGCAT	CTAATAGGTT	2160
TCGACGTTTT	TCTTTTTTTA	AGGTTTTTAT	TTGGTCTCCT	AGAATTTAAG	GTCTTAGTTA	2220
GTTTTGGTTT	GTTTTGTGGG	TTACATATTT	TCAATTCAAA	GGAGAATTTA	GCTGTCTTTT	2280
ATAATGTCCA	ATAGAGATAA	CGAGAGCATG	CTGCGTACTA	CATCAAGCGA	TAAGGCGATC	2340
GCTAGTCAAA	GGGATAAACG	GAAGTCTGAA	GTTTTTGATTG	CTGCACAGTC	CCTTGACAAAT	2400
GAAATCCGCA	GCGTAAAAAA	CCTAAAAAGA	TTGTCGATTG	GGTCAATGGA	TTTACTTATT	2460
GATCCAGAAT	TAGATATAAA	ATTCGGTGGG	GAATCTAGTG	GGAGACGATC	ATGGTCTGGC	2520
ACGACATCCA	GTTCTGCGTC	AATGCCAAGT	GACACAACCA	CCGTTAATAA	CACACGATAT	2580
AGCGATCCAA	CTCCGCTAGA	GAAC TTGCAT	GGGAGGGGTA	ACTCAGGGAT	AGAATCCTCC	2640
AATAAGACTA	AAATTAAATG	CTAACGTATT	AAAGAAAAAC	TTATTATGGG	TTCCCGCCAA	2700
TCAACACCCT	AACGTTAAGC	CTGATAATTT	CCTAGAGCTT	GTACAAGATA	CTTTACAAAA	2760
TATACAATA	AGCGACAATG	GTGAAGATAA	TGATGGGAAT	AGCAATGAAA	ATAACGATAT	2820
TGAGGATAAT	GGGGAGGATA	AAGAATCACA				2850

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 451 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURES:

- (A) NAME/KEY: peptide
- (B) LOCATION: 1...451
- (D) OTHER INFORMATION: /note = "The sequence of the Afclp protein from yeast presented as a polypeptide sequence"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Phe	Asp	Leu	Lys	Thr	Ile	Leu	Asp	His	Pro	Asn	Thr	Pro	Trp	Lys
1				5					10					15	
Leu	Ile	Ile	Ser	Gly	Phe	Ser	Ile	Ala	Gln	Phe	Ser	Phe	Glu	Ser	Tyr
			20					25					30		
Leu	Thr	Tyr	Arg	Gln	Tyr	Gln	Lys	Leu	Ser	Glu	Thr	Lys	Leu	Pro	Pro
		35				40						45			
Val	Leu	Glu	Asp	Glu	Ile	Asp	Asp	Glu	Thr	Phe	His	Lys	Ser	Arg	Asn
	50				55						60				
Tyr	Ser	Arg	Ala	Lys	Ala	Lys	Phe	Ser	Ile	Phe	Gly	Asp	Val	Tyr	Asn
	65				70				75					80	
Leu	Ala	Gln	Lys	Leu	Val	Phe	Ile	Lys	Tyr	Asp	Leu	Phe	Pro	Lys	Ile
			85					90						95	
Trp	His	Met	Ala	Val	Ser	Leu	Leu	Asn	Ala	Val	Leu	Pro	Val	Arg	Phe
			100					105						110	

His	Met	Val	Ser	Thr	Val	Ala	Gln	Ser	Leu	Cys	Phe	Leu	Gly	Leu	Leu
	115						120					125			
Ser	Ser	Leu	Ser	Thr	Leu	Val	Asp	Leu	Pro	Leu	Ser	Tyr	Tyr	Ser	His
	130					135					140				
Phe	Val	L	u	Glu	Glu	Lys	Phe	Gly	Phe	Asn	Lys	Leu	Thr	Val	Gln
145						150					155				160
Trp	Ile	Thr	Asp	Met	Ile	Lys	Ser	Leu	Thr	Leu	Ala	Tyr	Ala	Ile	Gly
				165					170					175	
Gly	Pro	Ile	Leu	Tyr	Leu	Phe	Leu	Lys	Ile	Phe	Asp	Lys	Phe	Pro	Thr
			180					185					190		
Asp	Phe	Leu	Trp	Tyr	Ile	Met	Val	Phe	Leu	Phe	Val	Val	Gln	Ile	Leu
	195						200				205				
Ala	Met	Thr	Ile	Ile	Pro	Val	Phe	Ile	Met	Pro	Met	Phe	Met	Lys	Phe
	210					215				220					
Thr	Pro	Leu	Glu	Asp	Gly	Glu	Leu	Lys	Lys	Ser	Ile	Glu	Ser	Leu	Ala
225					230					235					240
Asp	Arg	Val	Gly	Phe	Pro	Leu	Asp	Lys	Ile	Phe	Val	Ile	Asp	Gly	Ser
				245					250					255	
Lys	Arg	Ser	Ser	His	Ser	Asn	Ala	Tyr	Phe	Thr	Gly	Leu	Pro	Phe	Thr
				260				265					270		
Ser	Lys	Arg	Ile	Val	Leu	Phe	Asp	Thr	Leu	Val	Asn	Ser	Asn	Ser	Thr
	275						280					285			
Asp	Glu	Ile	Thr	Ala	Val	Leu	Ala	His	Glu	Ile	Gly	His	Trp	Gln	Lys
	290					295					300				
Met	His	Ile	Val	Asn	Met	Val	Ile	Phe	Ser	Gln	Leu	His	Thr	Phe	Leu
305					310					315					320
Ile	Phe	Ser	Leu	Phe	Thr	Ser	Ile	Tyr	Arg	Asn	Thr	Ser	Phe	Tyr	Asn
				325					330					335	
Thr	Phe	Gly	Phe	Phe	Leu	Glu	Lys	Ser	Thr	Gly	Ser	Ser	Phe	Val	Asp
			340					345					350		
Pro	Val	Ile	Thr	Lys	Glu	Phe	Pro	Ile	Ile	Ile	Gly	Phe	Met	Leu	Phe
	355						360					365			
Asn	Asp	Leu	Leu	Thr	Pro	Leu	Glu	Cys	Ala	Met	Gln	Phe	Val	Met	Ser
	370					375					380				
Leu	Ile	Ser	Arg	Thr	His	Glu	Tyr	Gln	Ala	Asp	Ala	Tyr	Ala	Lys	Lys
385					390					395					400
Leu	Gly	Tyr	Lys	Gln	Asn	Leu	Cys	Arg	Ala	Leu	Ile	Asp	Leu	Gln	Ile
				405					410					415	
Lys	Asn	Leu	Ser	Thr	Met	Asn	Val	Asp	Pro	Leu	Tyr	Ser	Ser	Tyr	His
			420					425					430		
Tyr	Ser	His	Pro	Thr	Leu	Ala	Glu	Arg	Leu	Thr	Ala	Leu	Asp	Tyr	Val
	435						440					445			
Ser	Glu	Lys													
	450														

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 315 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

(ii) MOLECULE TYPE: peptide

(ix) FEATURES:

- (A) NAME/KEY: peptide
- (B) LOCATION: 1...315

(D) OTHER INFORMATION: /note = "The sequence of the Rcelp protein from yeast presented as a polypeptide sequence"

Met	Leu	Gln	Phe	Ser	Thr	Phe	Leu	Val	Leu	Leu	Tyr	Ile	Ser	Ile	Ser	1	5	10	15
Tyr	Val	Leu	Pro	Leu	Tyr	Ala	Thr	Ser	Gln	Pro	Glu	Gln	Ser	Lys	Arg	20	25	30	
Asp	Asn	Pro	Arg	Thr	Ile	Lys	Ser	Arg	Met	Gln	Lys	Leu	Thr	Ile	Met	35	40	45	
Leu	Ile	Ser	Asn	Leu	Phe	Leu	Val	Pro	Phe	Leu	Gln	Ser	Gln	Leu	Ser	50	55	60	
Ser	Thr	Thr	Ser	His	Ile	Gly	Phe	Lys	Asp	Ala	Phe	Leu	Gly	Leu	Gly	65	70	75	80
Ile	Ile	Pro	Gly	Tyr	Tyr	Ala	Ala	Leu	Pro	Asn	Pro	Trp	Gln	Phe	Ser	85	90	95	
Gln	Phe	Val	Lys	Asp	Leu	Thr	Lys	Cys	Val	Ala	Met	Leu	Leu	Thr	Leu	100	105	110	
Tyr	Cys	Gly	Pro	Val	Leu	Asp	Phe	Val	Leu	Tyr	His	Leu	Leu	Asn	Pro	115	120	125	
Lys	Ser	Ser	Ile	Leu	Glu	Asp	Phe	Tyr	His	Glu	Phe	Leu	Asn	Ile	Trp	130	135	140	
Ser	Phe	Arg	Asn	Phe	Ile	Phe	Ala	Pro	Ile	Thr	Glu	Glu	Ile	Phe	Tyr	145	150	155	160
Thr	Ser	Met	Leu	Leu	Thr	Thr	Tyr	Leu	Asn	Leu	Ile	Pro	His	Ser	Gln	165	170	175	
Leu	Ser	Tyr	Gln	Gln	Leu	Phe	Trp	Gln	Pro	Ser	Leu	Phe	Phe	Gly	Leu	180	185	190	
Ala	His	Ala	His	His	Ala	Tyr	Glu	Glu	Leu	Gln	Glu	Gly	Ser	Met	Thr	195	200	205	
Thr	Val	Ser	Ile	Leu	Leu	Thr	Thr	Cys	Phe	Gln	Ile	Leu	Tyr	Thr	Thr	210	215	220	
Leu	Phe	Gly	Gly	Leu	Thr	Lys	Phe	Val	Phe	Val	Arg	Thr	Gly	Gly	Asn	225	230	235	240
Leu	Trp	Cys	Cys	Ile	Ile	Leu	His	Ala	Leu	Cys	Asn	Ile	Met	Gly	Phe	245	250	255	
Pro	Gly	Pro	Ser	Phe	Leu	Asn	Leu	His	Phe	Thr	Val	Val	Asp	Lys	Lys	260	265	270	
Ala	Gly	Arg	Ile	Ser	Lys	Leu	Val	Ser	Ile	Trp	Asn	Lys	Cys	Tyr	Phe	275	280	285	
Ala	Leu	Leu	Val	Leu	Gly	Leu	Ile	Ser	Leu	Lys	Asp	Thr	Leu	Gly	Thr	290	295	300	
Leu	Val	Gly	Thr	Pro	Gly	Tyr	Arg	Ile	Thr	Leu						305	310	315	